



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 1999

**The sense of nonsense lies in the nonsense of sense. Comment on Paolillo's
(1998) "Gary Larson's Far Side: Nonsense? Nonsense!" - Debate**

Ruch, Willibald

Abstract: The present work puts into perspective claims that have been made in a recent article by Paolillo (1998). It is pointed out that the conclusions drawn are based on misrepresentation and misunderstanding of our general approach as well as the aims of the criticized study by Kohler and Ruch (1994) in particular. Different ways of taxonomizing humor are discussed and it is concluded on the basis of both studies that Far Side humor falls well within the boundaries of the taxonomy of humor we developed.

DOI: <https://doi.org/10.1515/humr.1999.12.1.71>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-77514>

Journal Article

Published Version

Originally published at:

Ruch, Willibald (1999). The sense of nonsense lies in the nonsense of sense. Comment on Paolillo's (1998) "Gary Larson's Far Side: Nonsense? Nonsense!" - Debate. HUMOR: International Journal of Humor Research, 12(1):71-93.

DOI: <https://doi.org/10.1515/humr.1999.12.1.71>

The sense of nonsense lies in the nonsense of sense. Comment on Paolillo's (1998) "Gary Larson's *Far Side*: Nonsense? Nonsense!"

WILLIBALD RUCH

Abstract

The present work puts into perspective claims that have been made in a recent article by Paolillo (1998). It is pointed out that the conclusions drawn are based on misrepresentation and misunderstanding of our general approach as well as the aims of the criticized study by Köhler and Ruch (1994) in particular. Different ways of taxonomizing humor are discussed and it is concluded on the basis of both studies that Far Side humor falls well within the boundaries of the taxonomy of humor we developed.

Introduction

The conduct of cross-disciplinary collaboration in humor research has repeatedly been called for (e.g., Raskin 1995) and there have indeed been several symposia at ISHS conferences with participants from different disciplines and also at least one publication document where linguists and psychologists joined forces in a project (Ruch, Attardo, and Raskin 1993). To me the collaboration with linguists was inspiring, and it helped me to see the strengths and limitations of different disciplines including my own. However, I also remember that it took the three of us at least one afternoon at Brock University to understand each other's position, and to discuss the basic misunderstandings about each other's disciplines before one could come to constructive work.

Now John C. Paolillo enters the cross-disciplinary dialogue, and — as his work concerns not only psychological research in general but also my

approach — this should be welcome to me. However, it appears that little got clarified in the rather informal communication the two of us had when we met briefly at conferences, or exchanged letters. Since HUMOR recently published an article where Paolillo (1998) not only finds at least three main difficulties in a poster Gabriele Köhler and I presented at the 1994 ISHS conference in Ithaca, but also discovers serious oversights in and difficulties with my general work on a humor taxonomy (supplemented by a variety of recommendations for improvement), I think I should take the offer and answer and discuss his claims in public to reach the same audience his article did. In my reply I will point out that his claims and conclusions are wrong, and how (namely only by misunderstanding, misinterpreting, and misrepresenting my work) he could have arrived at them. First, though, I will briefly discuss different ways of thinking on the nature of the process of generating a taxonomy of humor and describe the taxonomy I used, and then address the elements of the controversy.

Taxonomizing of humor

How do we do it?

While there are many ways to construct a taxonomy of humor, I would like to draw attention to the different ways of approaching this goal at a more fundamental level. In detail, I would like to distinguish three different yet not exhaustive ways of thinking about the process of establishing a taxonomy of humor. One of these models assumes we only need to consider “humor” (Model A), one accepts that both “humor” and “researcher” have to be considered (Model B), and the third one considers “humor,” “researcher,” and “recipients” (Model C) as necessary. To not further complicate matters I am leaving out all other issues that would need to be considered when discussing taxonomies of humor.

Model A is the most narrowly defined approach to a taxonomy. It assumes that jokes/cartoons are different and the intrinsic structure only needs to be identified by somebody who takes the effort to do so (plus perhaps has the skills which some academic discipline would provide). The intrinsic structure will be identified independent of who takes the effort and once this endeavor has been undertaken, the problem is solved forever. This taxonomy would find approval by the academic community and would only be repeated once new humor has been generated and the need to test

the comprehensiveness arises again. Thus, the role of the *researcher* itself is not an explicit factor in that rationale, as it is not considered to determine the outcome (i.e., the final taxonomy). Consequently, for example, there is no need to have another person undertake the same steps to examine whether and to what extent his or her model will be different. Still, if two undertake the same steps and they disagree in their outcome they might start quarreling about who is right and who is wrong, but not acknowledge that they as persons are an integral part of the research endeavor and that different taxonomies can coexist.

Model B takes into account that different researchers have different backgrounds, different aims, and apply different tools, and thus also the resulting taxonomy will most likely be different. For example, depending on one's training and interests one might want to cluster for joke features such as repressed needs, historical or regional origin, linguistic structures, etc. It is evident that a taxonomy that was generated for one purpose will not necessarily be appropriate for other aims. Thus, Model B implies that there will not easily be a universal taxonomy that fits the needs of all researchers. Still one might form a team of experts from all disciplines and get a very general all-purpose taxonomy containing the totality of distinguishing criteria as accumulated by the different disciplines. However, even if all experts agree, do we have a taxonomy satisfying all needs? Maybe yes, but we would have an overkill of distinguishing features, most of them only relevant to those who analyze a joke and not to the everyday receiver of a joke.

Model C takes into account the naive receiver of humor as well. The importance of a joke's ingredients (i.e., what factors make jokes different or similar) would not exclusively be determined by the researcher's direct judgment of the joke but would be based on the perception of the respondents as well. Here we do not analyze jokes, or jokes as examined by the expert, but jokes as perceived by the layperson. We would place two jokes into the same category (or consider them to be similar) if they are perceived to be similar in a direct comparison, or if they independently from each other elicit comparable responses in the receiver. These two jokes might be different on many dimensions identified a priori by experts; if so, then apparently that identified dimension does not affect the naive person's response very much. Such an approach would state that if variations in a hypothetical joke parameter are below threshold for a recipient's response (i.e., they do not affect their response), then it is not high in importance for building a taxonomy of jokes. In model C the researcher is still important as there are several ways in which to incorporate the individual, what tools

to choose etc. Therefore, even psychologists — who are most likely to include the subjective element — will not necessarily come up with the same taxonomy when studying the same humor material and the same sample of individuals.

I have added these thoughts to describe the particularities of different approaches. Since in research articles there is rarely enough space to present all the implications of one's approach, this might be necessary for further discussion in general but also help illuminating why Paolillo's (1998) criticisms miss the point. In humor research as an interdisciplinary field there are different levels of analyses, different goals, different methodological backgrounds and hence, to me, it is not a problem if we do not agree in the way we taxonomize humor. Problems arise, though, once we start judging other approaches out of their context, e.g., whether or not they answer one's *own* questions.

The 3 WD is based on a Model C taxonomy of humor

The taxonomy of humor underlying the 3 WD humor test considers the subjects' cognitive-emotional responses to the stimuli. It is a taxonomy of humor as seen through the eyes of laypersons and systematized by a personality researcher. In short, we distinguish among the factor analytically derived humor categories of incongruity-resolution, nonsense, and sexual humor (*stimulus mode*), and we separate positive (funniness) and negative (aversiveness) components of humor appreciation (*response mode*). (For the development and validation of this two-mode model of humor appreciation see, for example, Ruch 1992).

How much of the taxonomy is determined by the three elements of Model C? As regards "humor" and "recipients," I have tested the robustness of the taxonomy against substitution of humor stimuli (i.e., I used different but overlapping sets of jokes and cartoons) and substitution of individuals (i.e., I replicated the taxonomy in different samples) and it appeared to be relatively stable. As regards the influence of the "researcher," the matter is more complicated. While a few people have analyzed the 3 WD stimuli in different samples (of even different nations) independent of me and roughly the same three-factor structure emerged (see Ruch and Hehl 1998), a stronger test of the role of the researcher is missing as nobody has undertaken a similar attempt starting all over, i.e., from different material. More importantly, it is easily demonstrated that my background in personality and my research goals predetermined some

early decisions to be made and consequently also affected the final taxonomy. For example, my choice of using factor analysis to taxonomize humor was determined by the goal to arrive at dimensions of humor, and these dimensions were thought to be eventually measured by a test that would ultimately be used to study the relationship between humor appreciation and personality based on an individual differences approach.

As these research goals and the agenda are not idiosyncratic, one might expect that other researchers sharing my background are likely to arrive at a similar or identical taxonomy. However, I do not think that this approach will suit all needs and it has limitations. For example, while its strength is to cluster humor according to the major dimensions (as perceived by the receiver), it is not really aimed at describing the *diversity* of humor within a factor. Just the opposite: in the 3 WD we use 10 to 15 jokes/cartoons per category and collapse the answers across them to form one total score. This is necessary to increase the reliability of the measurement but at the same time this means treating different jokes/cartoons to be exchangeable. While there are clear shortcomings with that approach, pursuing it has yielded a lot of interesting findings and insights into both appreciation of humor and personality characteristics. For example, we have gained the insight that structural factors are as important in clustering of humor as content (if not more so), and that humor appreciation is deeply rooted in one's personality (see, for example, Ruch 1992; Ruch and Hehl 1998). All in all, the 3 WD is not a "humor only" or "humor and researcher only" taxonomy and not keeping this apart inevitably induces confusion. It can be falsified using a Model C approach, but it does not make sense to evaluate it from a Model A framework.

The things to be clarified: Paolillo's claims

I will now turn to the misunderstandings and misrepresentation of the taxonomy that we developed and of the aims and outcomes of the Köhler and Ruch (1994) study.

Is nonsense humor NONsense: No!

The key element is Paolillo's distorted presentation of how we interpret our category of nonsense (NON) humor. The manuscript generates the

impression that our “nonsense” factor is exclusively or mainly based on jokes and cartoons where there is “no resolution.” This is repeatedly stated. For example, in the abstract Paolillo states that “... NON is characterized as having a structure in which there is no resolution,” and later he claims: “The best candidates for the 3 WD NON category are those with the resolution category NONE” (p. 273). Likewise, Paolillo capitalizes the “Non” in “nonsense” (i.e., NONsense), perhaps to imply that I believe that there is no resolution in nonsense humor, or that one possibly cannot understand the term in a different way than that. However, I never thought that the nonsense category consists exclusively of cartoons containing no resolution. This can be easily seen in the definition we gave (McGhee, Ruch and Hehl 1990: 124), where we elaborated that in nonsense humor “... the punch line may (1) provide no resolution at all, (2) provide a partial resolution (leaving an essential part of the incongruity unresolved), or (3) actually create new absurdities or incongruities.” We added that in nonsense humor the resolution information gives the appearance of making sense out of incongruities without actually doing so. Nonsense is a sophisticated play with our ability and tendency to make sense. We are misled to conclude things, resolve the incongruities, but then also realize that our tendency to make sense got misused and that the sense is actually nonsense. We referred to the model of Rothbart and Pien (1977) who proposed to distinguish between possible and impossible incongruities and between complete and incomplete resolutions. They argue that possible incongruities can be resolved completely while for impossible incongruity there is always a residue of incongruity left. This seemed to describe some of the cartoons in that factor well and we added these ideas.

When Paolillo quotes our definition of nonsense he correctly spells out that I also consider punch lines with incongruities that provide only a partial resolution, or whose resolution induces new incongruity as part of the nonsense category. At all other places he drops these elements and this allows him to come to the conclusions he draws. It is a curiosity, though, that he then himself raises the possibility “... that an assortment of resolution types are grouped together under NON” (p. 266) thereby neglecting that this is part of the definition of nonsense humor he keeps denying throughout the manuscript. Taking “nonsense” literally allows him to reckon “... lack of the appropriate background knowledge could reduce any humorous items to nonsense” (p. 281). He does not state that he *deliberately wants* to understand the factor of nonsense humor in a *different* way than I do. While I don’t mind if anybody wants to define nonsense

differently, it is also clear that I have to disagree when he later tries to suggest that the outcome of his study has implications for the present taxonomy¹. While Paolillo (1998: 272) first writes "... although the 3 WD treats partial resolution together with unresolved humor in the category NON, the prevalence of partial resolution in the *Far Side* ... made it seem advisable to treat this category separately for purposes of analysis," he then later forgets that he initially did so only "for purposes of analysis" and counts only NONE as a marker for nonsense. Only this strategy allows summarizing later that the hypothesis that *Far Side* cartoons represent nonsense (humor) is NONsense.

Does nonsense humor cover partial resolutions: Yes!

It may look deliberate but there is good reason to count partial resolution under nonsense. We discovered that some of the nonsense cartoons lying on the margins of the nonsense factor sometimes have also a low positive loading on INC-RES. Still when correlating these cartoons with personality variables, they behaved like the others in the NON category (and different, often with opposite signs, than those in the INC-RES category) and so it is obvious to leave them in the nonsense category. The common element is that they have residual incongruity and that apparently some people enjoy this. Thus, it seems that in this humor category it is not the resolution element that is important, but the degree of remaining incongruity.

This again illustrates the difference between a researcher-based and a recipient-based model. For a researcher PART may lie halfway between FULL and NONE, but laypersons might respond to PART more like they do to NONE and less so than they do to FULL. In *vitro* analysis of cartoons assumes that all features perceived by the analyst, and only these are perceived by and relevant for the recipient. However, in *vivo* experience as an experimenter told me something different. In my thesis I asked research participants to explain the jokes and to retell the events as a story. The outcome was that it is not uncommon that they project things into the joke which are not really there (or not easily seen by somebody else), emphasize peripheral things, and generally laugh about other things than intended. This clearly shows that one has to keep apart the differences between a Model A and Model C approach to humor. Does one want to classify joke texts according to features developed by a theorist, or jokes as

perceived by the layperson? Needless to say, *both* approaches are necessary and we can gain insight by applying them jointly.

Why study Far Side humor? The study by Köhler and Ruch (1994)

As Paolillo and others before point out, *Far Side* has been used frequently in research. There might be at least three reasons for studying *Far Side* humor. First, researchers may use the humor they like in their studies. This might be hazardous though, as sheer liking or admiration for a cartoonist is the wrong motivation for using his or her cartoons in an experimental study. Obviously, jokes are different, are liked by different people and have different effects, and not controlling for the type of humor used might result in unpredictable and nonreplicable results. Thus, it is more advisable to use the work of different cartoonists, and even better, different representatives of a comprehensive taxonomy of humor to balance out the different tastes of the research participants.

A second reason for choosing a particular cartoon series like the *Far Side* might be that these cartoons represent something special that is well suited for one's research purpose and which other humor does not provide. For example, Lefcourt and Shepherd (1995; Lefcourt 1996; see also Lefcourt, Davidson, Shepherd, and Phillips 1997) consider *Far Side* cartoons as a means to assess what they call *perspective taking humor*. Likewise, other cartoon series were considered to be well suited to represent certain topics and contents well (e.g., disgust humor, black humor). Relatedly, one may assume that a cartoon series really represents a *new* form of humor and this would motivate its use in studies. However, when pointing out these favorable particularities of a series or when considering them to be new and better suited for a particular aim, one should also provide empirical evidence that this is really the case. For example, one should try to relate it to existing similar humor, or better still, locate it in a comprehensive framework such as a taxonomy of humor. If *Far Side* humor is a hitherto undiscovered dimension of humor, then these cartoons would fall out of existing frameworks and would form an entirely new factor. Finally, a third aim might be the case study of an artist (or writing a biography). In this case one should be more interested in studying the *complete* works and not doing so would be a serious omission.

The poster by Köhler and Ruch (1994) was aimed at raising awareness for these problems, rather than pursuing any one of these aims. We argued that if one decides to use a newly emerged cartoon series, he or she should try to

locate this kind of humor in existing taxonomies of humor so that one can compare findings and accumulate knowledge in humor research. Thus, we did not see it to be our aim or responsibility to study *all Far Side* humor, or large samples of it, as we think this needs to be done by those who want to use *Far Side* humor and have their results be generalizable beyond it. For our own interest, to replicate the inclusion of Larson in one prior study (Forabosco and Ruch 1994), and as a service to those who wish to use *Far Side* cartoons and compare their findings with our nonsense humor category, we added eight cartoons to an ongoing study, which in total lasted several hours and therefore did not allow the inclusion of many more cartoons.

Is Far Side humor something new? Yes and no

Far Side humor is unique in many ways, such as in the format of the cartoon, some characters, the drawing style, etc. Most likely one could show people one *Far Side* cartoon from a distance of 10m and those familiar with *Far Side* humor would recognize its origin with a low error rate. Nevertheless, at a more global level *Far Side* humor is also comparable to what existed before. It serves a taste for humor that already existed. For example, the German cartoonists Karl-Friedrich Waechter (1978) and Robert Gernhardt (1983) utilized a somewhat similar humor style in some of their cartoon books and at least one of the cartoons is essentially identical. (In the *Far Side* version there is a scene with two polar bears, one of them about to eat an igloo including its human content, and saying to the other: "Oh hey! I just love these things! ... Crunchy on the outside and a chewy center!"; in the case of Waechter (1978: 8) it is two giants watching cars driving on a mountain road and one of them says: "The shells are hard, but the inside's real tender and tasty!!" [*Die Schale ist hart, aber innen sind sie weich und lecker.*]). In my factor analytic studies, cartoons of both artists marked the nonsense factor well. Thus, for several reasons, Köhler and Ruch (1994) doubted that *Far Side* humor is an entirely new dimension of humor, but expected that it would overlap with the 3 WD factor of nonsense humor, where all the cartoons by Gernhardt and Waechter load. To get a hunch whether the cartoons used for the assessment of perspective-taking humor are different from a randomly selected group of *Far Side* cartoons, we chose four cartoons for each of these two subcategories.

The size of the correlation between eight *Far Side* cartoons and the 15 standard cartoons of the NON category was .77 (Paolillo 1998); a

correlation approaching the reliability of the scales. The correlation for cartoons representing presumed perspective-taking humor and the other four cartoons were not any different ($r_s = .73$ in both cases, $p < .001$). Thus, we felt safe to conclude that "... while the *Far Side* cartoons do enrich the pool of nonsense cartoons, they do not challenge the comprehensiveness of the 3 WD taxonomy" (Köhler and Ruch 1994: 5). Thus, both our pilot study and Paolillo's analyses seem to confirm the major aim of the Köhler and Ruch study albeit in different ways, namely to demonstrate that *Far Side* can be located *within* the existing three dimensions and it is not necessary to expand the taxonomy. The multiple correlation between *Far Side* and all three humor categories is very high ($r = .79$, $d.f. = 3$ and 77 , $p < .001$). Likewise, in Paolillo's study of 800 cartoons 36 percent of the cartoons fall under NONE, 48 percent under PART (both corresponding to my nonsense factor) and the remaining 15.7 percent (of the category FULL) are described in a way that they might match the INC-RES factor, suggesting that even all 100 percent might be located within the 3 WD framework. So, if both studies did not make major mistakes, we can conclude that appreciation of *Far Side* humor is embedded into existing "tastes" of humor; thus, at a general level, there is no need to postulate a new dimension of humor appreciation not considered so far. (Likewise, our study did not show any difference between cartoons selected for the measure of perspective-taking and a random selection of *Far Side* cartoons). This interpretation is not in conflict with the view that — like many cartoonists — Gary Larson has his own unique style, and *Far Side* humor is clearly distinct at the surface level. Having answered this question, we move on to *where* in the system *Far Side* can best be located.

Is Far Side NONsense? No! Does Far Side fall under the nonsense humor factor? Yes, but ...!

It should be obvious by now that Köhler and Ruch did not claim that *Far Side* cartoons are NONsense (as in making no sense at all) as the 3 WD does not even foresee such a category. However, due to Paolillo's analysis of a high number of cartoons one can see that the number of *Far Side* cartoons that fit our nonsense category is quite high: 84.3 percent of the 800 cartoons were classified either under NONE or PART. Thus, while we were *not* interested in a comprehensive analysis of *Far Side*-humor, one can still say that our quasi-random small sample of cartoons did foresee the result of the more painstaking detailed analysis quite well and it seems safe to conclude

that investigators who want to search for markers of our NON category will find them easily among *Far Side* cartoons. Still, according to Paolillo's analysis not *all* cartoons seem to fall under nonsense.

What is the evidence that Far Side also goes under INC-RES?

There are two sources of evidence Paolillo uses. First, his reasoning is based on the significant correlations between aversiveness of the *Far Side* cartoons and the rated aversiveness of the 3 WD categories. This is not a good basis, though. As mentioned on different occasions (e.g., Ruch 1992), the aversiveness scores are always highly intercorrelated, and therefore the variance due to the categories is overlaid by a sort of general factor of aversiveness. Thus, it is not surprising that *Far Side* cartoons correlate also with the other categories and not only with nonsense. A partial correlation would be called for, but even without that, nonsense yields the highest correlations. Also, the distribution of ratings of aversiveness of individual cartoons is skewed and as INC-RES is found least aversive of all three categories, it is typically prone to yield a more skewed distribution. Furthermore, funniness is the major dimension in responses to humor, and for NON_f the pattern is very clear. As funniness scores are only mildly positively correlated it is acceptable also to find a low positive (but nonsignificant) correlation with INC-RES_f. Furthermore, in a stepwise regression analysis with funniness of the three humor categories as predictors and funniness of *Far Side* cartoons as criterion, only NON_f enters the equation and reduces the partial correlation coefficient for INC-RES_f to .07.

To study the locations of individual *Far Side* cartoons in the 3 WD factor space a reanalysis of the Köhler and Ruch (1994) data was undertaken. A joint principal components analysis of the 45 items of the 3 WD-K and the eight *Far Side* cartoons clearly yielded the three factors postulated by the 3 WD (Eigenvalues: 14.58, 6.18, 4.00, 2.15, 1.88, 1.73) which explained 48 percent of the variance. The obliquely rotated loadings of the *Far Side* cartoons are given in Table 1 along with the mean loadings of the items of all humor categories involved.

The factors were easily identified by the mean loading pattern of the 3 WD items: factor 1 loaded by the INC-RES items, factor 2 was marked by *both* the NON and the *Far Side* cartoons, and factor 3 represented SEX. Most importantly, the mean communality of the 3 WD items ($h^2 = .46$ and of the *Far Side* cartoons ($h^2 = .49$) did not differ; thus, the factors accounted

Table 1. Location of the Far Side cartoons and the 3 WD categories in a joint factor space (reference structure)

	Factor 1	Factor 2	Factor 3
<i>Cartoons</i>			
Far Side 1 [#]	.04	.48	.10
Far Side 2*	-.05	.69	.15
Far Side 3 [#]	.27	.50	.05
Far Side 4*	.02	.53	.06
Far Side 5 [#]	.00	.73	-.08
Far Side 6*	-.10	.71	-.08
Far Side 7*	-.12	.60	.22
Far Side 8 [#]	-.06	.69	.13
<i>Mean loadings</i>			
Far Side cartoons (<i>n</i> = 8)	-.03	.65	.08
INC-RES (<i>n</i> = 15)	.60	.01	.08
NON (<i>n</i> = 15)	.04	.56	.04
SEX (<i>n</i> = 15)	.02	.05	.55

Notes: INC-RES = Incongruity-Resolution humor, NON = Nonsense humor, SEX = Sexual humor (all from the 3 WD).

[#]randomly selected *Far Side* cartoons, **Far Side* cartoons representing perspective-taking humor (Lefcourt *et al.* 1995) (*N* = 81).

for about half of the variance of the two sets of humor items, again clearly confirming that the *Far Side* cartoons are *within* the boundaries of the present taxonomy. Furthermore, Table 1 shows that *all eight Far Side* cartoons clearly belong to the nonsense factor. As the reference structure controls for the intercorrelations among the factors, the matrix allows for a better estimation of loadings of the *Far Side* cartoons on the other factors. While no second loading exceeded .30, item 3 clearly had an alignment with the INC-RES factor (factor 1). However, this is not unusual as also the items of the 3 WD test may be factorially complex, and all in all, the factorial complexity of the *Far Side* cartoons (1.17) was even (numerically, but not statistically) *lower* than the one of the 3 WD (1.33) items.

The second evidence is Paolillo's finding that 16 percent of the *Far Side* cartoons he analyzed fell into his FULL category. This is a high number and it should not be neglected. Still, one might argue that considering how strongly humor and personality are connected, it seems to be unlikely that somebody who provides such fine markers for one category will be also producing cartoons that go along with almost opposite personality characteristics. Thus, it might be of interest to examine how the cartoons Paolillo classified as FULL would load in our system (whether they clearly

load on INC-RES or are located more in between INC-RES and NON) if subjected to a factor analysis together with our material. However, such a study does not have high priority for me and I do not have any particular academic interest in *Far Side* cartoons whatsoever, because the copyright restrictions that Paolillo and others report make it unlikely that *Far Side* cartoons will be ever included in a research instrument. Of course, if it turns out that these cartoons would clearly fall outside the INC-RES and NON dimensions, this would be a challenge and one would need to start thinking whether the system needs to be expanded.

Paolillo's findings: What is new is not true, and what is true is not new

I will not comment on flaws in Paolillo's article and procedure in general as this is the task of a review process. However, a few comments are necessary as they have impact on the present discussion. From what I will point out it will become clear that Paolillo's study does not allow for the claims he makes.

First, Paolillo applies a theory meant for verbal humor (the *General Theory of Verbal Humor* by Attardo and Raskin 1991) to *graphical* material. It is not important whether the six knowledge resources also play a role in cartoons; the question is what *additionally* needs to be considered when studying graphical material. Here Paolillo does not give the artist his due as he does not acknowledge that the drawing style contributes much to Gary Larson's uniqueness. This assumption could easily be tested in an experiment, where a set of original *Far Side* cartoons are compared (according to relevant criteria, such as perceived funniness) with the same set of cartoons redrawn by a layperson, or with their "joke"-version (i.e., a verbal recount of the cartoon that considers all the elements considered important by Paolillo). Even without performing this experiment we can see how much a cartoon loses in funniness, by looking at Paolillo's description of his example cartoons. As Paolillo's "in-depth examination" pretends to analyze *all* the relevant features in *Far Side* humor, his work has to be judged according to this claim, and I doubt that his analysis involves all the relevant features that make *Far Side* cartoons unique. While the study by Köhler and Ruch (1994) does so even less, this was *not* our aim.

Furthermore, there is a problem regarding the conduct of the study as he seems to pursue a Model A approach. More specifically, Paolillo (1998) does not give any information on *objectivity* and *reliability* of the coding

procedure. He himself determined the number and nature of the categories, and assigned all the cartoons to these categories. Both predictor and criteria variables were judged by the *same* person allowing for all sorts of confounding. No part of the whole procedure was done independently by another person which would allow for comparing the outcome and determining the degree of intersubjective agreement. It might well be that the results Paolillo presents might just reflect his idiosyncratic view of these cartoons.

The flaw in the design and the inappropriate use of significance tests are equally fatal as they prohibit generalizations. Paolillo analyzed *all* cartoons of four books representing the *early* period of the work, not a random sample drawn from a defined large population. He overlooks that the sample studied equals the population about which one would like to make inferences, and this makes statistical tests as the Chi-Square analyses of tables 2 to 4 *obsolete* and hence *invalidates* the conclusions drawn (e.g., "... We now have evidence of significant statistical relationships between resolution type and three independent variables ..." Paolillo 1998: 277). When one tests the population, inference is neither needed nor appropriate, and the frequency values can be compared directly without a test of "significance." In this sense then, a score of, say 5341.10 is higher than one of 5341.09; i.e., any numerical difference counts. Furthermore, one can also *not make statistical generalizations* to other populations; i.e., other forms of humor, or humor *per se*. The validity of Paolillo's findings is restricted to these 800 cartoons and, strictly speaking, cannot even be generalized to the later period of Gary Larson, much less allowing him to say that his "... closer inspection of the *Far Side*" does "shed some light" (p. 266) on the nature of the factors of the 3 WD. If he had in mind to make generalizations, he should have defined a population of humor and then drawn a random sample of cartoons. If his results should be relevant for the 3 WD, the sample of humor items needed to come from the same population.

Finally, I have to question the novelty of what Paolillo actually found. The relationship between his real and unreal situations and the type of resolution he claims to have found is nothing new. Actually, it can be seen as a reformulation of the model suggested by Rothbart and Pien (1977) which proposed to distinguish between possible and impossible incongruities and between complete and incomplete resolutions. They argue that only possible incongruities can be resolved completely while for an impossible incongruity only a partial resolution is possible and a residue of incongruity is left. Paolillo was apparently unaware of this model as he quotes this

source as “Rothbart and Pien (1977, cited in Raskin 1985: 33)”. The rediscovery of the model described by Rothbart and Pien in *Far Side* cartoons gives further credibility to the hypothesis that *Far Side* cartoons can be subsumed under the nonsense humor factor, as I, in part, found this model appropriate to describe some of the cartoons of this factor (see above).

On claims and untested hypotheses

When discussing why my findings could be different than they are, Paolillo (1998) raised several hypotheses but he did not test their validity. For example, he writes “A third possibility is that the *Far Side* cartoons ... are interpreted out of context by the German raters” (p. 266). This hypothesis is already raised in the introduction and reiterated throughout the manuscript, almost like a fact (“It is shown that ... their German raters are likely to lack crucial aspects of the cultural knowledge of TA and LM involved in interpreting *Far Side* cartoon,” p. 261) but no attempts to test this hypothesis are made. I think it is fine to raise a hypothesis in the discussion and treat it as a possibility. Or, if one has the hypothesis *before* the study, one should design an appropriate study to test the hypothesis and, if it gets confirmed, write an article about it. But here the claim is repeated throughout the text without adding any information that might help in deciding whether it is true or not. While I will not comment on the plausibility of this hypothesis, I think those who believe in this hypothesis should test it and *then* report whether it applies or not.

Likewise, the examples of potential problems with the factor analytic approach given by Paolillo are well considered and theoretically correct but they do not match with reported data I know. To my experience there are no “really good jokes” which are universally appreciated by all potential raters (see the data in Ruch and Hehl 1998). Therefore, I cannot count this argument.

On raising and not clarifying problems

Paolillo makes several assumptions about what we have done or not done, and what we perhaps might have done wrong, but he does not attempt to clarify the problems he raises although this partly would be easy to do. For example, he writes “No 3WD studies undertake structural analyses of the test stimuli, nor are any particular exemplars of the three types described in

the studies I have seen (see references)” (Paolillo 1998: 266). In his references the studies of Ruch and/or Köhler in which the 3 WD is used totals 4(!) articles, suggesting that only about 10–15 percent of the available published sources have been examined. Given this low number, the above claims look quite bold.

Paolillo’s work has been impaired as we — due to the space constraints given by the format of a poster presentation — did not reprint the eight cartoons and he writes: “. . . Köhler and Ruch (1994) did not describe the *Far Side* cartoons they used or identify the pages that they might be found on, otherwise it would be possible to discuss how representative those cartoons are of the *Far Side* in general” (Paolillo 1998: 273). I only can add that if he had asked for them we would certainly have been glad to send them to him, as we did send him a copy of our poster presentation. Also, we did identify half of the cartoons by referring that we use the same ones that have been used in a publication on *perspective-taking humor* that we quoted in the poster.

Likewise, while I can accept that not seeing the 3 WD makes work difficult (“Without direct access to the test forms, or at least some examples, it is difficult to evaluate the interpretation of the 3 WD humor categories,” p. 266), I have to add that we are not hiding examples of the 3 WD factors, or the 3 WD test in general. I typically do not provide examples in articles for several reasons; (a) as Paolillo experienced himself, there is a copyright problem, (b) there are space restrictions and describing cartoons and jokes takes up space, (c) it is good academic tradition to exchange letters regarding discussion of research findings, questions raised by them or only asking for clarification or materials used. So typically there is the sentence “Address correspondence and request for reprints to the author at . . .” at the end of my articles. Several colleagues have taken advantage of this offer and so far I have sent inspection copies to everybody who has asked for them, and likewise I asked for and always received the materials colleagues used in their studies. Furthermore, I have reprinted examples in articles in both languages I publish in; examples of the cartoons can even be found in one article I wrote for HUMOR. Thus, as we have been in mail contact, most of these problems that hindered Paolillo’s work could have been solved easily by simply writing to one of us.

What exactly is wrong with the 3 WD studies?

In the section “Conclusions and Cautions” Paolillo (1998) raises concerns about the approach I took, speculates about what I might have done wrong,

and also gives recommendations how to remedy the purported weak current state. The problems mentioned refer to the tool of factor analysis, that we have not tested the stability of the taxonomy against substitution of social groups, and that the interpretation was done without considering alternative interpretations. Thus, I need to discuss the extent to which these claims are true.

As regards factor analysis as a research tool there have been many more problems raised throughout the history of that approach. In fact, criticisms of this approach are almost as old as the method itself, as this tool was sometimes not used with the appropriate circumspection. Some of the criticisms have led to the refinement of the tool and a more careful application of it, some proved to be irrelevant and few limitations still remain and need to be considered. Still, the advantages outweigh the disadvantages by far and make this a prime research tool in personality research. In my articles I don't repeat these criticisms as this would be highly redundant and this is common knowledge among those familiar with the method. I think that I have applied factor analysis with appropriate circumspection but I will certainly correct flaws once they are pointed out to me in more detail.

As regards the stability of a factor structure, Paolillo (1998) is right in pointing out this is an important issue. However, I have to clearly reject Paolillo's claim that I excluded "*a priori*" the possibility that the factor structure might differ among social groups or cultural background. Paolillo (1998: 284) claims that "... hypotheses concerning the possible variation of the patterns across social groups are not considered" and more specifically "it is also a serious oversight, where such salient social categories such as gender and age are involved." I can add in my defense that especially the analyses of the initial larger samples involved the testing of the stability of the factor structure in subgroups involving gender, age and other variables (the list of which should be mentioned in the respective articles). Also, we inspected the factor structure when we later tested specific samples, such as cardiac heart disease patients. I have extended the test of the taxonomy across different nations, such research, of course, being always contingent on the existence of researchers of these countries interested in collaboration (for the results, see Ruch and Hehl 1998). Paolillo is right in pointing out in his footnote that so far we did not extend this test to African and Australian aborigines, but if we did, he might still rightfully raise the question: what about astronauts? Or undercover agents? Thus, in principle this is a never ending story as *per nature* one will never be able to test it against *all* possible subgroups. Therefore, instability of factor structure will be found

inevitably, as the sheer number of tests necessary produces differences just by chance. Facing this impossible (and rather boring) task I surrender and leave it to others to demonstrate that the factor structure is moderated by a particular variable that I did not control for.

Paolillo criticized that in the cross-cultural studies the "... sample of stimuli is fixed, small, and not necessarily native to the cultures they are tested in" (p. 288). The number of items tested is 60 (i.e., 20 per humor factor), which is large enough to obtain a stable factor structure. Furthermore, if one wants to compare factor patterns, the number of items has to be fixed, and using native humor for both samples would impair if not prohibit any comparison. Finally, Paolillo overlooked that the articles comparing the taxonomy across nations did contain a sentence, such as: "while it can be claimed that the intrinsic structure in the 3 WD-humor pool is stable across the European countries studied so far, these results do not imply that there are no additional humor categories in those countries or in other countries not tested. Joint factor analyses of the 3 WD item pool *and* humor material selected to represent potential new humor categories should be carried out to answer this question."

A final caution that Paolillo (1998: 285) raises concerns the interpretation of the factors of the 3 WD in as much they "... do not necessarily admit of only one interpretation. There are enough different psychological theories of humor, some in which resolution is regarded as necessary, others in which it is not, that an interpretation of statistically isolated factors in terms of resolution types should be supported carefully, and alternative characterizations, both social and psychological, should be considered." Here Paolillo gets the history wrong. At the time I did my taxonomic studies (Ruch 1981, 1984), typically in research on personality and humor appreciation the classification of jokes was done according to joke *content*, not structure. The only prior study considering formal variables in classification to my knowledge was Eysenck's (1942) classic study. Both incongruity and incongruity-resolution theory existed, but as competing theories aimed at explaining the *same* material (i.e., they were claiming the status of universal models of humor), not as separate mechanisms that might be differently valid for different types of jokes. Also, they were seen as "givens" in humor, not affecting the perceived funniness of jokes *differentially*; and thus not needing to be considered in taxonomies. Thus, the preclassification of the items I studied was done according to the variables that were discussed up until that time, including, for example, the Freudian trilogy. With the exception of sexual humor — which was easily identified as all sexual

jokes and only those loaded on it — none of the discussed variables seemed to suit as a common denominator for the two remaining factors (as both, for example, had cartoons preclassified as being “aggressive”). Only after I realized that the two stage-model of humor appreciation proposed by Suls (1972) was working well mainly for factor one and that the resolutions provided by the cartoons of the second factor were less complete and better described by the model of Rothbart and Pien (1977), I decided to shift the emphasis from the content away to structural features in trying to answer what items of a category have in common. Thus, contrary to Paolillo’s assumptions, I was fully aware of the theories of that time and indeed I can even assert that my studies have shifted the direction of personality based humor research away from considering only content to studying also structural properties. Ironically, not many have made this move and current studies still try to predict content variables without controlling for the major variance — which I think is structure-related.

One has to distinguish different aims in describing humor categories. As stated elsewhere, I was not interested in a detailed description of the variety of the jokes of a factor, but in what they have *in common*. This will certainly leave out, for example, how *Far Side* cartoons are different from the ones by Waechter and Gernhardt. Furthermore, as with related constructs, the procedure in this research tradition is to outline a model, use that to derive hypotheses and test them in studies, and use that information to eventually refine the initial model. As the predictions based on the present model were in agreement with the outcomes there was no need to revise the model so far. This does *not* mean that nothing more can or should be said about the psychological make-up of the factors.

How to improve the taxonomy?

While so far the 3 WD taxonomy has been fruitful for research on the relationship between humor and personality, like most models, it sooner or later will be obsolete and get replaced by a better one. I would be disappointed if this were not the case as this would mean that it is not even worth bothering. In a recent chapter Ruch and Hehl (1998) summarized some findings on the 3 WD and gave further directions on where the model might need to go both in terms of refining the structural model (e.g., testing for further, perhaps more specific content categories) and the interpretation of the factors. I will not repeat the proposals made here but I certainly would like to use this occasion to add that the former would ideally be done

cross-nationally and the latter could certainly include contributions from other disciplines including linguistics.² While levels of analyses are different, they are certainly not incompatible. What it takes is the will and venturesomeness for interdisciplinary collaboration.

Conclusions

Paolillo does not seem to understand or want to accept that there are different approaches to a taxonomy and that in an *empirically* derived approach the definition of the dimensions is secondary, whereas in a *rationally* derived taxonomy the definition of the dimensions comes first. He also seems to believe that everything that an armchair theorist might isolate from analyses of cartoons, and only these features, are of actual importance to the recipient. Starting from misunderstanding and misrepresenting our approach he blames Köhler and Ruch for not reaching goals we did not intend to pursue, and generates the impression that I conducted a poor research program. He builds up a straw man by distorting how I understand nonsense humor and consequently tries to destroy this straw man by bringing his deliberately changed understanding of nonsense into conflict with what he claims to have found. His study, however, is discovered to be full of fundamental flaws largely prohibiting any interpretation. He does not acknowledge that even his analysis confirms that the conclusions of Köhler and Ruch were right: the *Far Side* cartoons do fit well into the boundaries of the 3 WD, even if perhaps not *all* cartoons fall under nonsense and perhaps incongruity-resolution needs to be considered as well. Most of the arguments I listed in the manuscript will not be new for Paolillo as I have already outlined them to him on one occasion or the other. However, they are important for the readers of HUMOR to judge for themselves whether his claims are warranted or unwarranted.

I have to add that I also disapprove of the whole procedure. Paolillo reprinted a data table from a poster without asking for permission, making a later publication of a more extensive report difficult. Also, a poster should be regarded as a prepublication; it is restricted in length (ours was 5 pages, including abstract and tables) and cannot go into detail as a final manuscript can. In our case the data collection was not even completed and the final sample size was increased albeit only slightly. He should have waited until our results are formally published and then start writing something.

Apart from these issues, I have to say that I do not understand the rationale of Paolillo's research program. If he wants to analyze *Far Side*

cartoons he should focus on them in even more detail and tell us what he learned about *Far Side* humor. If he wants to analyze the 3 WD he may do so and tell us what he found about the taxonomy from his point of view. But I do not see the logic in analyzing 800 *Far Side* cartoons to make comments on the taxonomy underlying the 3 WD. I also don't understand Paolillo's frame of reference in his judgments. It was about 20 years ago that I started work on a taxonomy of jokes and cartoons and I have spent some time over the years validating and expanding this approach in different directions. While some colleagues contributed to the research agenda and collaborated with me, unfortunately not many other research programs were aimed at arriving at a taxonomy of humor during this time. Thus, we still are far away from one of the first aims in a field, namely to provide a framework that allows one to integrate research findings so that accumulation of knowledge is possible. Now Paolillo comes and judges my work on goals I had not intended to reach, sets criteria which no single person or research team can achieve, and blames me for not pursuing questions I was not interested in. If he wants to contribute meaningfully to the present research agenda then he could have continued the dialogue we started and entered a fruitful collaboration. Indeed I can still see several ways in which interdisciplinary research in humor is needed. However, in order for this to work out, it is necessary that researchers talk to each other and try to understand the strengths and limitations of the different disciplines and where they supplement each other. In summary, I hope that interdisciplinary research in humor continues and perhaps our debate helps to clarify the different positions and to discipline determined starting points.

University of Düsseldorf

WILLIBALD RUCH

Notes

Correspondence and requests for reprints should be sent to Willibald Ruch, University of Düsseldorf, Department of Physiological Psychology, Universitätsstraße 1, 40225 Düsseldorf, Germany (e-mail: wruch@uni-duesseldorf.de). This manuscript was written during my stay at the University of Western Ontario. I would like to thank Rod Martin for polishing up the English and for helpful comments on an earlier version of this manuscript.

1. Actually it was Robert Gernhardt's (a cartoonist whose cartoons marked the nonsense factor well) description of his work as a form of "subtle nonsense" which played a role in deciding to label the second factor as "nonsense" humor. When using this label I did not

think that anybody would ever misunderstand the meaning of the factor as something that makes absolutely no sense, is incomprehensible, or simply some unconnected elements that have no relationship. In fact it was immediately clear that nonsense is a "higher" form of humor, more artistic, a very sophisticated play with our ability and tendency to make sense. "Residual incongruity" could have been a good alternative term but as high scores in nonsense humor also correlated with the questionnaire assessment of "liking nonsense" (Ruch 1980), this label did seem to fit well. (Perhaps it helps to remember that construct labels do not necessarily need to contain all the information the definitions provide, or even no information, as in the concept labels "Type A" or R. B. Cattell's "Harria" and "Premsia"). The perceived superior quality of nonsense humor has guided the search for certain personality variables that were expected to predict appreciation of nonsense, and the list of predictors found clearly confirms that this hunch was right (for predictors of humor appreciation, see Ruch 1992; Ruch and Hehl 1998).

2. Actually, I had extended the offer to analyze the 3 WD to two linguists but a collaboration on such a project did not materialize for different reasons.

References

- Attardo, Salvatore and Victor Raskin
 1991 Script theory revis(it)ed: Joke similarity and joke representation model. *Humor* 4: 347–411.
- Eysenck, Hans-Jürgen
 1942 The appreciation of humor: An experimental and theoretical study. *British Journal of Psychology* 32: 295–309.
- Forabosco, Giovannantonio and Willibald Ruch
 1994 Sensation seeking, social attitudes, and humor appreciation in Italy. *Personality and Individual Differences* 16: 515–528.
- Gernhardt, Robert
 1983 *Gernhardts Erzählungen. 120 Bildgeschichten von Robert Gernhardt*. Zürich: Hoffmanns Verlag.
- Köhler, Gabriele and Willibald Ruch
 1994 Testing the comprehensiveness of the 3 WD taxonomy of humor: Does Gary Larson's *Far Side* Gallery fit in? *International Society for Humor Studies Conference*, Ithaca College, NY, June 22–26, 1994.
- Lefcourt, Herbert M.
 1996 Perspective-taking humor and authoritarianism as predictors of anthropocentrism. *Humor* 9: 57–71.
- Lefcourt, Herbert M. and Robert S. Shepherd
 1995 Organ donation, authoritarianism, and perspective-taking humor. *Journal of Research in Personality* 29: 121–138.
- Lefcourt, Herbert M., Karina Davidson, Robert Shepherd and Margory Phillips
 1997 Who likes "Far Side" humor? *Humor* 10: 439–452.
- McGhee, Paul E., Willibald Ruch and Franz-Josef Hehl
 1990 A personality-based model of humor development during adulthood. *Humor* 3: 119–146.
- Paolillo, John C.
 1995 Gary Larson's "*Far Side*" — What kind of nonsense is this? In George E. C. Paton and Robert W. Jones (eds.), *1995 13th International Humor Conference*. Birmingham, England: University of Aston: 40.

- 1998 Gary Larson's *Far Side*: Nonsense? Nonsense! *Humor* 11: 261–290.
- Raskin, Victor
- 1985 *Semantic Mechanisms of Humor*. Dordrecht: Reidel.
- 1995 The state of the art in humour research: Let's amalgamate! In George E. C. Paton and Robert W. Jones (eds.), *1995 13th International Humor Conference*. Birmingham, England: University of Aston: 44.
- Rothbart, Mary K. and Diana Pien
- 1977 Elephants and marshmallows: A theoretical synthesis of incongruity-resolution and arousal theories of humour. In Anthony J. Chapman and Hugh C. Foot (eds.), *It's a Funny Thing, Humour*. Oxford: Pergamon Press: 37–40.
- Ruch, Willibald
- 1981 Witzbeurteilung und Persönlichkeit: eine trimodale Analyse. *Zeitschrift für Differentielle und Diagnostische Psychologie* 2: 253–273.
- 1984 Konservativismus und Witzbeurteilung: Konvergenz gegenstandsber-eichsinterner und übergreifender Variabilität? *Zeitschrift für Differentielle und Diagnostische Psychologie* 5: 221–245.
- 1991 An appeal for a cross-cultural approach to a taxonomy of humor. In Willibald Ruch and Victor Raskin (chair), *Approaches to the Taxonomy of Humor. Methodological Issues and Preliminary Results*. Symposium, Ninth International Conference on Humour and Laughter. St. Catherines, Canada, June 26–30, 1991.
- 1992 Assessment of appreciation of humor: Studies with the 3 WD humor test. In: Charles D. Spielberger and James N. Butcher (eds.), *Advances in Personality Assessment (Vol. 9)*. Hillsdale, NJ: Lawrence Erlbaum Associates, 27–75.
- Ruch, Willibald, Salvatore Attardo and Victor Raskin
- 1993 Towards an empirical verification of the General Theory of Verbal Humor. *Humor* 6:123–136.
- Ruch, Willibald and Franz-Josef Hehl
- 1998 A two-mode model of humor appreciation and its relationship with aesthetic sensitivity. In: Willibald Ruch (ed.), *The Sense of Humor: Explorations of a Personality Characteristic*. Humor Research Series, vol. 3. Berlin: Mouton de Gruyter, 109–142.
- Suls, Jerry M.
- 1972 A two-stage model for the appreciation of jokes and cartoons: An information-processing analysis. In Jeffrey H. Goldstein and Paul E. McGhee (eds.), *The Psychology of Humor*. New York: Academic Press, 81–100.
- Waechter, Friedrich-Karl
- 1978 *Wahrscheinlich guckt wieder kein Schwein*. Zürich: Diogenes Verlag.

